FLOATING GATE NON-VOLATILE MEMORY CELL WITH LOW ERASING VOLTAGE AND MANUFACTURING METHOD

Field of the Invention

This invention relates to a low erase
voltage, floating gate non-volatile memory cell. The
invention also concerns a method of making a low erase

5 voltage memory cell. Specifically, the invention
relates to a non-volatile memory cell of the type which
comprises at least one floating gate transistor
realized over a semiconductor substrate, with a source
region and a drain region separated by a channel region

10 which, in turn, is overlaid by a thin layer of gate
oxide to isolate from the substrate a floating gate
region coupled to a control gate terminal. The
invention relates, particularly, but not exclusively,
to a memory cell comprising a single transistor and

15 intended for storage devices of the FLASH type, and the
description herein below deals with this particular
application for convenience of illustration only.

Background of the Invention

As is well known, memory cells having a

20 single floating gate transistor are most widely
utilized in FLASH memories. To better illustrate the
invention aspects, a conventional type of memory cell,
as schematically shown in Figure 1, will be considered
first. Such a cell 1 comprises a MOS transistor 2

25 which is realized over a semiconductor substrate 3, and

has a source region S, and a drain region D separated by a channel region C which, in turn, has a control gate terminal GC realized thereover. This is, therefore, a transistor of the NMOS type wherein a
5   second gate GF, referred to as the floating gate on account of it being fully enveloped in a dielectric material, is also provided between the control gate GC and the substrate 3. In particular, the floating gate GF is separated from the control gate GC by a layer 4,
10   referred to as the interpoly dielectric layer. A thin layer 5 of silicon oxide, referred to as the gate oxide layer, is interposed between the floating gate GF and the substrate 3.

Thus, the floating gate GF cannot be accessed
15   from outside, and its potential will only depend on any charge which may be present thereon and on the capacitive couplings with the other elements of the cell. Accordingly, such a cell forms basically a data storage element based on charges becoming trapped on
20   the floating gate GF. Major advantages of a floating gate memory cell are: its compact size, low sensitivity to disturbances during the programming step, and low voltage requirements during the reading step, that is, for both the programming operations and
25   the cell erasing.

Such a memory cell 1 is usually programmed by hot electron injection and erased by a tunneling effect. The programming operation of a memory cell such as the cell 1 uses the generation of channel hot
30   electrons in the drain region D (or rather, in an area of spatial charge thereof known as the pinch-off area). Through the application of a favorable electric field, i.e., a field which enables a certain number of electrons of the drain region D to acquire sufficient
35   energy to overcome the potential barrier between the drain region and the gate oxide, charges are caused to

migrate toward the floating gate GF. The floating gate GF will store an essentially negative charge.

In a dual mode, for the erasing operation, which allows electrons to be extracted from the
5 floating gate GF of the memory cell 1, the tunneling effect is utilized, also referred to as Fowler-Nordheim's effect. In particular, current programming techniques provide for the erasing to be carried out by applying a high voltage to the source region S, this
10 operation being commonly referred to as source erasing. Under these conditions, the shape of the source junction should be a gradual one, so that the substrate current can be lowered and the risk of creating hot holes reduced.

15 The need for a gradual source junction shape represents, however, one of the strictest limitations imposed on the scalability of the gate length of a memory cell so realized.

A first known approach for overcoming this limitation
20 is to perform the operation of erasing of the memory cell 1 by a so-called channel erasing. In this case, the memory cell 1 is erased by applying a negative voltage to the control gate GC of the transistor 2, while holding the source S and drain D regions at the
25 same potential as the semiconductor substrate 3.

Channel erasing mitigates the limitations on the length scalability of the cell 1, thereby removing the need for a particular shape of the source junction. While being in many ways advantageous, this first
30 discussed approach has a major drawback in that it involves the application of high erase voltages to the control gate GC, due to the high capacitive coupling between the channel and the floating gate. In fact, while for source erasing an erase voltage of 10 to 12
35 volts is adequate, channel erasing requires a voltage of 17 to 20 volts. Such a high voltage value is,

however, incompatible with advanced CMOS processes for manufacturing semiconductor devices.

To address this problem, it has been prior practice to locate the storage device in a diffusion
5 well of the P-well type, isolated from the substrate 3, using a triple-well CMOS process. In this way, positive voltages can be applied to the substrate 3 and the source and drain regions of the memory cell during the erasing operation, so as to split the voltage between
10 the negatively biased control gate GC and the positively biased substrate 3. This would, accordingly, restrict the absolute value of the erase voltage needed.

With a view to obtaining a scalability for
15 the length of the cell 1 it is necessary, however, to concurrently raise the inherent threshold voltage V_{t0} of the memory cell 1 so as to attenuate the effects of a possible shorting of the channel region C (punch-through, drain-turn-on). The inherent threshold voltage
20 V_{t0} of the memory cell depends on the characteristics of the cell construction (such as the thickness of the gate oxide 5 or the capacitive ratios) and the dopant concentration in the channel region C.

The erase voltage required to have the cell
25 erased within a predetermined time, usually set by the storage device specifications, will be higher, the higher the inherent threshold voltage V_{t0} of the cell. In fact, the threshold voltage V_t of the cell depends on the charge Q stored in the floating gate GF, and the
30 inherent threshold voltage V_{t0} of the cell (that is, when it has no charge in the floating gate), according to the following relationship:

$$V_t = V_{t0} - Q/C_{pp}$$

being the coupling capacitance between the floating
35 gate GF and the control gate GC.

In addition, for the memory reading operations to be carried out properly, it is necessary for the threshold voltages of the memory cells to acquire suitable values, regardless of their inherent threshold

5 voltage Vt0.
In particular, memory cells containing a logic state "1" should have a threshold voltage typically within the range of 0.5V to 3V, whereas memory cells containing a logic state "0" should have a threshold

10 voltage typically above 5V.

As for memory cells of the FLASH type, current manufacturing processes provide for inherent threshold voltages between 2V and 3V. Thus, erased cells (i.e. cells with a logic state "1") will have a threshold

15 voltage Vt close to their inherent threshold voltage Vt0, or in other words, have a reduced number of charges on the floating gate GF, whereas programmed cells (i.e. cells with a logic state "0") will show a threshold voltage Vt which is definitely higher than

20 their inherent threshold voltage Vt0. In this context, the erasing operation includes extracting the negative charge stored in the floating gate GF of the programmed cells until it is restored to a condition of near-neutral charge.

25 In addition, the presence of charge in large amounts in the floating gate of the programmed cells contributes to an increased electric field in the gate oxide during erasing. As mentioned above, to obtain a scalability for the length of the cell 1, the dopant

30 concentration in the channel region C must be increased so as to attenuate the punch-through and drain-turn-on effects. In this way, memory cells can be obtained which are scaled with inherent threshold voltages Vt0 of 4 to 6 volts.

35 However, such cells are in a near-neutral charge condition in the programmed state of "0", and have a high positive charge in the floating gate GF when in

the erased state. This positive charge lowers the
electric field in the channel oxide during the erasing
operation and enforces the application of higher erase
voltages. Thus it occurs that, while a negative erase
5   voltage of 17-20V must be applied to the gate of a cell
having an inherent threshold voltage Vt0 of 2-3V, in
order to erase a memory cell with an inherent threshold
voltage of 4-6V, a negative erase voltage of 20-24V
must be provided. The need for scaling the cell length
10 would, therefore, lead to the use of high erase
voltages, which clashes with advanced CMOS processes
that have a limited capability to handle high voltages.

Summary of the Invention

In view of the foregoing background it is
15 therefore an object of the invention to provide a
scalable memory cell of the FLASH type, in particular
one comprising a single floating gate transistor, which
has such construction and performance characteristics
to permit a reduced erase voltage.
20   The idea of solution on which the instant
invention stands is to provide a memory cell which has
an asymmetrical potential barrier between the
substrate, gate oxide and floating gate region, to
thereby facilitate the erasing operation without
25 impairing the retention of data into the memory. Based
on this idea of solution, the technical problem is
solved by a memory cells as previously indicated and
defined in the characterizing portion of Claim 1. The
technical problem is also solved by a method of
30 manufacturing a memory cell as previously indicated and
defined in the characterizing portion of Claim 10.

Brief Description of the Drawings

The features and advantages of a memory cell
according to the invention will be apparent from the
35 following description of embodiments thereof, given by

way of example and not of limitation with reference to
the accompanying drawings. In the drawings:

        . Figure 1 shows schematically a single-transistor
memory cell according to the prior art;

5   Figure 2 shows schematically symmetrical potential
barriers between the substrate/gate oxide/floating gate
junctions of the cell in Figure 1;

        Figure 3 shows schematically a single-transistor
memory cell according to this invention;

10   Figure 4 shows schematically asymmetrical
potential barriers between the substrate/gate
oxide/floating gate junctions of a first embodiment of
the memory cell according to this invention;

        Figure 5 shows erase band patterns for
15 conventional memory cells and for a second embodiment
of the memory cell according to this invention;

        Figures 6A, 6B and 6C show schematically the
floating gate surface of different examples of the
second embodiment of the memory cell according to this
20 invention; and

        . Figure 7 shows schematically asymmetrical
potential barriers between the substrate/gate
oxide/floating gate junctions of a third embodiment of
the memory cell according to this invention.

25   Detailed Description of the Preferred Embodiments
        Referring to the drawing figures, generally and
schematically shown at 10 is a memory cell according to
the invention. It should be emphasized that, whereas
non-scalable cells according to the prior art have a
30 logic state "1" which is characterized by neutral
charge, and a logic state "0" characterized by the
presence of negative charge on the floating gate GF,
with conventional cells of the scalable type, the logic
state "1" corresponds to the presence of positive
35 charge on the floating gate GF, and the logic state "0"
corresponds to a near-neutral charge condition. The

gate oxide 5 of these cells should allow charge to be flowed by a high electric field being applied during the erasing operation, while ensuring retention of data stored in the memory. In other words, the gate oxide 5

5 of a memory cell, such as the cell 1 shown in Figure 1, should exhibit a very small leakage current in the presence of reduced electric fields, that is, during the reading operation or in a stand-by condition. In the latter case, the electric field would correspond to

10 the internal electric field developed by the charge stored in the floating gate.

Shown schematically in Figure 2 are first 6 and second 7 potential barriers occurring in the memory cell 1 according to the prior art. In particular, there

15 may occur two different situations, as specified below. With a conventional non-scalable cell type, wherein the charge stored on the floating gate is almost nil (in the logic state "1") or even negative (in the logic state "0"), the retention of data in the memory during

20 an erasing operation is ensured by the first potential barrier 6 provided between the substrate 3 and the gate oxide 5. This first potential barrier 6, in fact, will resist the injection of electrons into the floating gate GF of the cell during the reading step. This

25 phenomenon is referred to as charge gain.

Likewise in the stand-by condition, this retention in the memory is ensured by the second potential barrier 7 provided between the floating gate GF and the gate oxide 5. The second barrier 7 will resist the

30 leakage of electrons from the floating gate GF. This phenomenon is referred to as charge loss, by analogy with the former. Under such conditions, it is convenient if a memory cell is provided which has the first 6 and second 7 potential barriers substantially

35 identical with each other.

To achieve this, a semiconductor/gate oxide/floating gate structure should be arranged to

have a substantially symmetrical band pattern. Such a
cell has the same resistance to the electrons flow from
the substrate 3 to the floating gate GF as from the
floating gate GF to the substrate 3. It should be
5 considered that this concerns the electrons, not the
holes, since the contribution of holes to the gate
current is negligible. On the other hand, with cells
of the scalable type, for which the charge stored on
the floating gate GF is almost nil (in the logic state
10 "0") or even positive (in the logic state "1"), it is
only the first potential barrier 6 between the
substrate 3 and the gate oxide 5 that plays a deciding
role for a correct retention of data stored in the
memory during the reading operations or in the stand-by
15 condition.

The memory cell 10 of this invention is a cell
comprising a single floating gate transistor, which is
structurally similar to the conventional cell 1 of
Figure 1. This cell 10, shown schematically in Figure
20 3, comprises a floating gate transistor 12 realized
over a semiconductor substrate 13 with a source region
S1 and a drain region D1 separated by a channel region
C1 overlaid by a thin layer 15 of gate oxide which
isolates, from the substrate 13, a floating gate region
25 GF1 coupled to a control gate terminal GC1.

In the light of the above considerations, a method
is proposed of manufacturing a memory cell 10 of the
low erase voltage non-volatile type, whereby a first
potential barrier 16 is provided between the
30 semiconductor substrate 13 and the layer 15 of gate
oxide, and a second different potential barrier 17 is
provided between the floating gate region GF1 and the
layer 15 of gate oxide. In particular, the method of
this invention provides for the second potential
35 barrier 17 to be lower than the first potential barrier
16, thereby permitting the use of a lower erase voltage
Ver and ensuring the retention of any data stored in

the memory cell 10. Modified embodiments of the memory cell 10 according to the invention will now be described which have asymmetrical potential barriers between the substrate/gate oxide/floating gate

5   junctions. In particular, examples of memory cells having asymmetrical potential barriers will be described, which are based essentially on three embodiments of the invention.

Embodiment 1

10   Shown in Figure 4 are the first 16 and second 17 potential barriers, being asymmetrical relative to each other and associated with the memory cell 10 of this invention. A first manner of obtaining such first 16 and second 17 potential barriers is based on a

15   selection of a suitable material for forming the floating gate region GF1. In prior memory cells, the material used has been polycrystalline silicon which, having the same electron affinity as the silicon included in the substrate 3, provides a first 6 and a

20   second 7 potential barrier, identical with each other, as explained in connection with Figure 2.

   To provide asymmetrical barriers 16 and 17, a material must be used instead which has a lower electron affinity (or work function, in the instance of

25   metals) than silicon. Particularly effective, for the purpose of limiting the erase voltage Ver to be applied to the memory cell 10, are materials whose electron affinity lies between 2.5eV and 3.5eV, and which can provide potential barriers in excess of 1.7eV, i.e.,

30   appropriate to secure the data stored in the memory, while reducing the electric field required for the erasing operations. Advantageously, this first embodiment of the memory cell 10, according to the invention, is essentially characterized by a very

35   simple implementation.

Embodiment 2

A second way of implementing the memory cell 10
showing asymmetry in the conduction of current through
the gate oxide provides for the electric field to be
5   focused onto the surface of the floating gate GF1.
This focusing can be obtained through the so-called
point effect. Figure 5 shows an erase band pattern SB1
of the memory cell 1 according to the prior art, and an
erase band pattern SB2 for a second embodiment of the
10   inventive memory cell 10. It also illustrates
schematically the electric field focusing effect for
the band pattern SB2.

The operation of the second embodiment of the
memory cell 10 according to the invention will now be
15   explained. As a result of the focusing effect, the
erase voltage Ver is increased due to the increased
local electric field. The same result may also be
obtained with floating gates implemented by
polycrystalline silicon, the height of the barriers of
20   potential remaining the same as in the prior art.

In particular, the electric field focusing effect
can be obtained by physically providing a plurality of
·points 18 on the surface of the floating gate GF1 in
the direction toward the gate oxide 15. A first method
25   for obtaining this plurality of points 18 comprises
chemi-physical techniques of etching away the gate
oxide layer 15 prior to depositing the polycrystalline
silicon, as shown schematically in Figure 6A.   A
second method provides instead for the deposition of a
30   thin dielectric layer 19 onto the surface of the gate
oxide layer 15, thereby to bring about a nucleation
effective to break the continuity of the dielectric
layer whereover the polycrystalline silicon layer will
then be deposited, as shown schematically in Figure 6B.
35   Finally, a third method includes introducing
granules 20 of a conductive material into the gate
oxide layer 15, in the vicinity of the floating gate

region GF1, as shown schematically in Figure 6C. These granules may be introduced, for example, by a low-energy implantation of silicon (or another suitable material) in the gate oxide layer 15, followed by a
5 thermal treatment of the memory cell 10 to precipitate the implanted silicon, thereby forming the granules 20. Alternatively, a thin layer of polycrystalline silicon could be deposited, followed by oxidation, such that residual silicon granules not fully oxidized can later
10 be isolated using, of preference, a grain-edge oxidation process.

Embodiment 3

A third way of implementing the memory cell 10 having asymmetrical potential barriers, according to
15 the invention, is to deposit a double layer as the gate dielectric. The double layer includes a first layer 5 of silicon oxide at the substrate 13 surface, and a second layer 21 of a suitable dielectric material whose electron affinity lies intermediate to those of the
20 silicon oxide and the polycrystalline silicon of the floating gate GF1, as shown schematically in Figure 7. Advantageously, two or more of the above methods can be used in combination, with this invention.

In a preferred embodiment of the memory cell 10 of
25 this invention, the gate oxide layer 15 is realized using silicon nitride. Silicon nitride is a dielectric material which has a higher electron affinity than silicon oxide (a prerequisite of the third embodiment); in addition, it lends itself for use in combination
30 with a layer of silicon oxide. In this case, it is desirable that the layer of silicon nitride be a relatively thin one, to avoid undesired effects of charge build-up at the interface between the silicon oxide and nitride. Also known is that a deposition of
35 silicon nitride at a low temperature results in nucleation of silicon precipitates within the nitride

matrix. Such deposition would, therefore, enable the
formation of focusing spots as explained in connection
with the second embodiment of the invention. In this
case, a combination of the second and the third
5   embodiment would be provided.